

WHAT IS CLAIMED IS:

1. A method of determining sequence variations in a target biomolecule, comprising:

- 5 a) cleaving the target biomolecule into fragments by contacting the target biomolecule with one or more specific cleavage reagents;
- b) cleaving or simulating cleavage of a reference biomolecule into fragments with the same cleavage reagent(s);
- c) determining mass signals of the fragments produced in a) and b);
- 10 d) determining differences in the mass signals between the fragments produced in a) and the fragments produced in b); and
- e) determining a reduced set of sequence variation candidates from the differences in the mass signals and thereby determining sequence variations in the target compared to the reference biomolecule.

15 2. The method of claim 1, wherein the biomolecule is a biopolymer.

 3. The method of claim 1, wherein the biomolecule is polypeptide.

 4. The method of claim 1, wherein the biomolecule is a nucleic acid.

20 5. The method of claim 1, wherein the biomolecule is DNA.

 6. The method of claim 1, wherein the biomolecule is RNA.

 7. The method of claim 1, further comprising scoring the candidate sequences and determining the sequence variations in the target nucleic acid molecule.

8. The method of claim 1, wherein determining a set of reduced sequence variation candidates comprises:

a) identifying fragments that are different between the target biomolecule and the reference biomolecule;

5 b) determining compomers corresponding to the different fragments identified in step a) that are compomer witnesses; and

c) determining a reduced set of sequence variations corresponding to the compomer witnesses that are candidate sequences to determine the sequence variations in the target compared to the reference

10 biomolecule.

9. The method of claim 1, wherein the differences in mass signals are manifested as missing signals, additional signals, signals that are different in intensity, and/or as having a different signal-to-noise ratio.

10. The method of claim 1, wherein the masses are determined
15 by mass spectrometry.

11. A method of determining sequence variations in a target nucleic acid molecule, comprising:

a) cleaving the target nucleic acid molecule into fragments by contacting the target nucleic acid molecule with one or more specific
20 cleavage reagents;

b) cleaving or simulating cleavage of a reference nucleic acid molecule into fragments using the same cleavage reagent(s);

c) determining mass signals of the fragments produced in a) and b);

d) determining differences in the mass signals between the
25 fragments produced in a) and the fragments produced in b); and

e) determining a reduced set of sequence variation candidates from the differences in the mass signals and thereby determining sequence variations in the target compared to the reference nucleic acid.

12. The method of claim 11, wherein determining a set of reduced
5 sequence variation candidates comprises:

a) identifying fragments that are different between the target nucleic acid and the reference nucleic acid;

b) determining compomers corresponding to the identified different fragments in step a) that are compomer witnesses; and

10 c) determining a reduced set of sequence variations corresponding to the compomer witnesses that are candidate sequences to determine the sequence variations in the target nucleic acid compared to the reference nucleic acid.

13. The method of claim 11, wherein the differences in output
15 signals are manifested as missing signals, additional signals, signals that are different in intensity, and/or as having a different signal-to-noise ratio.

14. The method of claim 11, wherein the masses are determined by mass spectrometry.

15. The method of claim 11, wherein the sequence variation
20 is a mutation or a polymorphism.

16. The method of claim 11, wherein the mutation is an insertion, a deletion or a substitution.

17. The method of claim 15, wherein the polymorphism is a single nucleotide polymorphism.

18. The method of claim 1, wherein the target is a target nucleic acid molecule from an organism selected from the group consisting of eukaryotes, prokaryotes and viruses.

19. The method of claim 18, wherein the organism is a bacterium.

5 20. The method of claim 19, wherein the bacterium is selected from the group consisting of *Helicobacter pyloris*, *Borelia burgdorferi*, *Legionella pneumophila*, *Mycobacteria sp.* (e.g. *M. tuberculosis*, *M. avium*, *M. intracellulare*, *M. kansaii*, *M. gordonae*), *Staphylococcus aureus*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Listeria*
 10 *monocytogenes*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus sp.*, *Streptococcus faecalis*, *Streptococcus bovis*, *Streptococcus pneumoniae*, *Campylobacter sp.*, *Enterococcus sp.*, *Haemophilus influenzae*, *Bacillus anthracis*, *Corynebacterium diphtheriae*, *Corynebacterium sp.*, *Erysipelothrix rhusiopathiae*, *Clostridium*
 15 *perfringens*, *Clostridium tetani*, *Enterobacter aerogenes*, *Klebsiella pneumoniae*, *Pasturella multocida*, *Bacteroides sp.*, *Fusobacterium nucleatum*, *Streptobacillus moniliformis*, *Treponema pallidum*, *Treponema pertenue*, *Leptospira* and *Actinomyces israeli*.

20 21. The method of claim 1, wherein a specific cleavage reagent is an RNase.

22. The method of claim 21, wherein the specific cleavage reagents are selected from among the RNase T₁, RNase U₂, the RNase PhyM, RNase A, chicken liver RNase (RNase CL3) and cusavitin.

25 23. The method of claim 1, wherein a specific cleavage reagent is a glycosylase.

24. The method of claim 1, wherein sequence variations in the target biomolecule permit genotyping a subject, forensic analysis, disease diagnosis or disease prognosis.

25. The method of claim 11, wherein the method determines epigenetic changes in a target nucleic acid molecule relative to a reference nucleic acid molecule.

26. The method of claim 11 that is a method for determining allelic
5 frequency in a sample, comprising:

a) cleaving a mixture of target nucleic acid molecules in the sample containing a mixture of wild-type and mutant alleles into fragments using one or more specific cleavage reagents;

b) cleaving a nucleic acid molecule containing a wild-type allele into
10 fragments using the same cleavage reagent(s);

c) determining mass signals of the fragments;

d) identifying fragments that are different between the mixture of target nucleic acid molecules and the wild-type nucleic acid molecule;

e) determining compomers corresponding to the identified different
15 fragments in step d) that are compomer witnesses;

f) determining allelic variants that are candidate alleles corresponding to each compomer witness;

g) scoring the candidate alleles; and

h) determining the allelic frequency of the mutant alleles in the
20 sample.

27. The method of claim 26, wherein the allelic frequency is about 5-10%

28. The method of claim 26, wherein the allelic frequency is less than 5%

29. A method for determining sequence variations at one or more base positions in a plurality of target nucleic acid molecules, comprising:

a) cleaving the target nucleic acid molecules into fragments by contacting the molecules with one or more specific cleavage reagents;

5 b) cleaving or simulating cleavage of one or more reference nucleic acid molecules into fragments using the same cleavage reagents;

c) determining the mass signals of fragments produced a) and b);

d) identifying fragments that are different between the target nucleic acid molecules and the one or more reference nucleic acid
10 molecules;

e) determining compomers corresponding to the different fragments that are compomer witnesses;

f) determining the sequence variations that are candidate sequences corresponding to each compomer witness;

15 g) scoring the candidate sequences; and

h) determining the sequence variations in the plurality of target nucleic acid molecules.

30. The method of claim 29, wherein after cleaving the target nucleic acid molecules and the one or more reference molecules into
20 fragments, the fragments are immobilized on a solid support.

31. The method of claim 30, wherein the fragments comprise an array.

32. The method of claim 29, wherein the specific cleavage reagents are selected from among RNase T₁, RNase U₂, RNase PhyM,
25 RNase A, chicken liver RNase (RNase CL3) and cusavitin.

33. The method of claim 29, wherein a specific cleavage reagent is a glycosylase.

34. The method of claim 30, wherein the array is a chip for mass spectrometry.

5 35. A method for detecting sequence variations in a target nucleic acid in a mixture of nucleic acids in a sample, comprising:

a) performing more than one specific cleavage reaction using the same or different specific cleavage reagents on the sample, wherein the target nucleic acid is cleaved in a plurality of fragmentation reactions to

10 generate a plurality of fragmentation patterns;

b) performing or simulating more than one specific cleavage reaction on a reference nucleic acid under conditions that are the same as those of the target cleavage reactions in step a);

c) determining the fragments that are different between the
15 plurality of fragmentation patterns of the cleaved target nucleic acid and the plurality of fragmentation patterns of the cleaved reference nucleic acid;

d) determining the different fragments that are consistent with a particular sequence variation in the target nucleic acid;

20 e) combining the consistent different fragments corresponding to one or more sequence variations to obtain a spectrum of different fragments;

f) determining, from the spectrum of different fragments, those different fragments containing compomers that are compomer witnesses;

25 g) determining the sequence variations that are candidate sequences corresponding to each compomer witness;

h) scoring the candidate sequences; and

i) determining the sequence variations in the target nucleic acid molecule in a mixture of nucleic acids in a biological sample.

36. The method of claim 35, wherein the biological sample is a
5 tumor sample.

37. The method of claim 35, wherein the biological sample comprises genomic DNA from a pool of individuals.

38. The method of claim 36, wherein about 5-10% of the mixture of target nucleic acids contains the sequence variations.

10 39. The method of claim 36, wherein less than 5% of the mixture of target nucleic acids contains the sequence variations.

40. A program product for use in a computer that executes program instructions recorded in a computer-readable media to determine sequence variations in a target biomolecule, the program product
15 comprising:

a recordable media; and

a plurality of computer-readable program instructions on the recordable media that are executable by the computer to perform a method comprising:

20 a) determining mass signals of target biomolecule fragments produced from cleaving a target biomolecule into fragments by contacting the target biomolecule with one or more specific cleavage reagents and determining mass signals of a reference biomolecule fragments produced from cleaving or simulating cleavage of a reference biomolecule into
25 fragments using the same cleavage reagents;

b) determining differences in the mass signals between the fragments produced in the target biomolecule and the fragments produced in the reference biomolecule; and

- c) determining a reduced set of sequence variation
5 candidates from the differences in the mass signals and thereby determining sequence variations in the target compared to the reference biomolecule.

41. The program product of claim 40, wherein the computer executable method further comprises scoring the candidate sequences
10 and determining the sequence variations in the target biomolecule.

42. The program product of claim 40, wherein determining a set of reduced sequence variations of the computer executable method further comprises:

a) identifying fragments that are different between the target
15 biomolecule and the reference biomolecule;

b) determining compomers corresponding to the different fragments identified in step a) that are compomer witnesses; and

c) determining a reduced set of sequence variations corresponding to the compomer witnesses that are candidate sequences
20 to determine the sequence variations in the target compared to the reference biomolecule.

43. The program product of claim 40, wherein the differences in output signals are manifested as missing signals, additional signals, signals that are different in intensity, and/or as having a different signal-
25 to-noise ratio.

44. The program product of claim 40, wherein the masses are determined by mass spectrometry.

45. A program product for use in a computer that executes program instructions recorded in a computer-readable media to determine sequence variations in a target nucleic acid molecule, the program product comprising:

5 a recordable media; and

a plurality of computer-readable program instructions on the recordable media that are executable by the computer to perform a method comprising:

- a) determining mass signals of target nucleic acid molecule
- 10 fragments produced from cleaving a target nucleic acid molecule into fragments by contacting the target nucleic acid molecule with one or more specific cleavage reagents and determining mass signals of a reference nucleic acid molecule fragments produced from cleaving or simulating cleavage of a reference nucleic acid molecule into fragments
- 15 using the same cleavage reagents;
- b) determining differences in the mass signals between the fragments produced in the target nucleic acid and the fragments produced in the reference nucleic acid; and
- c) determining a reduced set of sequence variation
- 20 candidates from the differences in the mass signals and thereby determining sequence variations in the target compared to the reference nucleic acid.

46. The program product of claim 45, wherein the masses of the fragments are determined by mass spectrometry.

25 47. A combination of the program product of claim 40 and one or more specific cleavage reagents.

48. A system, comprising a computer, the program product of claim 40, and one or more specific cleavage reagents.

49. The combination of claim 47, further comprising:
one or more reference nucleic acid molecules; and/or one or more natural
5 or modified nucleoside triphosphates.

50. A kit for determining sequence variations in one or more target nucleic acid molecules, comprising a combination of claim 47 and optionally instructions for determining sequence variations.

51. The kit of claim 50, wherein a specific cleavage reagent is an
10 RNase.

52. The kit of claim 51, wherein the RNases are selected from among the RNase T₁, RNase U₂, the RNase PhyM, RNase A, chicken liver RNase (RNase CL3) and cusavitin.

53. A computer-based method for identifying sequence variations
15 in a target nucleic acid molecule or plurality thereof, comprising:

a) entering a reference sequence and the identity of one or more specific cleavage reagents into the computer;

b) entering the masses of the fragments generated by reaction of the same cleavage reagent(s) with a target nucleic acid molecule;

20 c) identifying fragments that are different between the target nucleic acid molecule and the reference nucleic acid molecule;

d) determining compomers corresponding to the identified different fragments in step c) that are compomer witnesses;

e) determining the sequence variations that are candidate
25 sequences corresponding to each compomer witness;

g) scoring the candidate sequences; and

h) determining the sequence variations in the target nucleic acid molecule or a plurality thereof.

54. The method of claim 53, wherein in step d), the compomers
5 that are compomer witnesses are determined by comparing the compomers corresponding to the identified different fragments generated in step c) to a database of previously determined compomer witnesses for each of the specific cleavage reagents.

55. A system for high throughput analysis of sequence variations
10 in a target nucleic acid molecule in a sample, comprising:

a processing station that performs a fragmentation reaction, in the presence of one or more specific cleavage reagents, on a target nucleic acid molecule in a reaction mixture;

a robotic system that transports the resulting fragmentation
15 products from the processing station to a mass measuring station, wherein the masses of the products of the reaction are determined; and

a data analysis system that processes the data from the mass measuring station by performing the method of claim 53 to identify sequence variations at one or more positions in the target nucleic acid
20 molecule in the sample.

56. The system of claim 55, further comprising a control system that determines when processing at each station is complete and, in response, moves the sample to the next test station, and continuously processes samples one after another until the control system receives a
25 stop instruction.

57. The system of claim 55, wherein the mass measuring station is a mass spectrometer.

58. The method of claim 11, wherein prior to cleaving the target nucleic acid molecule into fragments, the nucleic acid is treated so that the cleavage specificity is altered.

59. A method for determining single nucleotide polymorphisms at one or more base positions in a plurality of target nucleic acid molecules, comprising:

a) cleaving the target nucleic acid molecules into fragments by contacting the molecules with one or more base specific cleavage reagents;

b) cleaving or simulating cleavage of one or more reference nucleic acid molecules into fragments using the same cleavage reagents;

c) determining the mass signals of fragments produced in a) and b);

d) identifying fragments that are different between the target nucleic acid molecules and the one or more reference nucleic acid molecules;

e) determining compomers corresponding to the identified different fragments in step d) that are compomer witnesses;

f) determining single nucleotide polymorphisms in candidate sequences corresponding to each compomer witness;

g) scoring the candidate sequences; and

h) determining the single nucleotide polymorphisms in the plurality of target nucleic acid molecules.

60. The method of claim 59, wherein the specific cleavage reagent is an RNase.

61. The method of claim 59, wherein the specific cleavage reagents are selected from among the RNase T₁, RNase U₂, the RNase PhyM, RNase A, chicken liver RNase (RNase CL3) and cusavitin.

62. The method of claim 59, wherein the target nucleic acids
5 molecules are selected from among single stranded DNA, double stranded DNA, cDNA, single stranded RNA, double stranded RNA, DNA/RNA hybrid, PNA (peptide nucleic acid) and a DNA/RNA mosaic nucleic acids.

63. The method of claim 59, wherein the target nucleic acids are produced by transcription.

10 64. The method of claim 59, wherein the target nucleic acids comprise genomic DNA from a pool of individuals.

65. A method of determining single nucleotide polymorphisms in a target nucleic acid molecule, comprising:

a) cleaving the target nucleic acid molecule into fragments by
15 contacting the target nucleic acid molecule with one or more base specific cleavage reagents;

b) cleaving or simulating cleavage of a reference nucleic acid molecule into fragments using the same cleavage reagent(s);

c) determining mass signals of the fragments produced in a) and b);

20 d) determining differences in the mass signals between the fragments produced in a) and the fragments produced in b); and

e) determining a reduced set of single nucleotide polymorphism candidates from the differences in the mass signals and thereby determining single nucleotide polymorphisms in the target compared to
25 the reference nucleic acid.

66. The method of claim 65, wherein the specific cleavage reagent is an RNase.
67. The method of claim 65, wherein a specific cleavage reagents are selected from among the RNase T₁, RNase U₂, the RNase PhyM, 5 RNase A, chicken liver RNase (RNase CL3) and cusavitin.
68. The method of claim 65, wherein the target nucleic acids molecule is selected from among single stranded DNA, double stranded DNA, cDNA, single stranded RNA, double stranded RNA, DNA/RNA hybrid, PNA (peptide nucleic acid) and a DNA/RNA mosaic nucleic acid.
- 10 69. The method of claim 65, wherein the target nucleic acid is produced by transcription.
70. The method of claim 65, wherein the target nucleic acid is genomic DNA from a single individual.
71. The method of claim 65, futher comprising scoring the 15 reduced set of single nucleotide polymorphism candidates.
72. The method of claim 65, further comprising scoring heterozygous single nucleotide polymorphism candidates.
73. The method of claim 65, further comprising scoring homozygous single nucleotide polymorphism candidates.